THE SAMPLE AUTOCORRELATIONS OF HEAVY-TAILED PROCESSES WITH APPLICATIONS TO ARCH[1]

By Richard A. Davis and Thomas Mikosch

Colorado State University and University of Groningen

We study the sample ACVF and ACF of a general stationary sequence under a weak mixing condition and in the case that the marginal distributions are regularly varying. This includes linear and bilinear processes with regularly varying noise and ARCH processes, their squares and absolute values. We show that the distributional limits of the sample ACF can be random, provided that the variance of the marginal distribution is infinite and the process is nonlinear. This is in contrast to infinite variance linear processes. If the process has a finite second but infinite fourth moment, then the sample ACF is consistent with scaling rates that grow at a slower rate than the standard $\sqrt{n}$. Consequently, asymptotic confidence bands are wider than those constructed in the classical theory. We demonstrate the theory in full detail for an ARCH(1) process.

1. Introduction. The motivation for this paper comes from an empirical observation which has been made in the econometrics and applied financial time series literature for a long time: log-returns of various series of share prices, stock indices, exchange rates and interest rates are believed to be heavy-tailed in the sense that their fourth moments can be infinite. A second observation is that such time series appear to have a complicated dependence structure that cannot be adequately modeled with a linear process. Departures from linearity can often be detected by examination of the sample autocorrelation function, (ACFs) of a time series, their absolute values and squares. For log-returns, the sample ACF of the original series vanishes at almost all lags (with possible significant values at the first two or three lags) while the sample ACF of the absolute values and squares are usually nonzero, and decay to zero slowly. Since this behavior is not consistent with that expected from a linear process, a variety of nonlinear models has been proposed for such data. The ARCH (autoregressive conditionally heteroscedastic) process and its extensions and generalizations form one of the more popular and extensively studied classes of nonlinear time series models, which exhibit many of the properties described above. In this paper, we will

use the ARCH(1) process to illustrate our limit results for the sample ACF and sample autocovariance function (ACVF) of a stationary time series.

It is the aim of this paper to give a general theory for the sample ACF and sample ACVF of a stationary sequence $(X_t)_{t \in \mathbb{Z}}$. Two assumptions are made: there exist a mixing condition weaker than strong mixing and a regular variation condition on the finite-dimensional distributions of the process. Under these assumptions, the point process theory in Davis and Hsing [8] is extended to the case of multivariate stationary processes $(\mathbf{X}_t)$. In particular, the weak convergence of the point processes $N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1} \mathbf{X}_t}$ for appropriate normalizing constants a_n is established. This result is then applied to the stationary process constituted by the random vectors $a_n^{-1}(X_t, \ldots, X_{t+m})$. As a next step, weak convergence of (N_n) is used to obtain the weak convergence of the point processes based on the products $a_n^{-2}(X_t^2, X_t X_{t+1}, \ldots, X_t X_{t+m})$. Finally, the points $a_n^{-2} X_t X_{t+h}$ of the lagged process are summed up, and so the joint convergence of the sample autocovariances and autocorrelations at different lags is established.

Our main results on the asymptotic behavior of the sample ACVF and ACF in the infinite second and fourth moment cases are contained in Theorem 3.5. The limits of the suitably normalized sample ACVF and ACF are expressed in terms of the limiting point processes. Limit distributions are ratios of infinite variance stable vectors or functions of such vectors. Hence tabulation of the quantiles from the limit distributions is difficult. In the case of infinite variance X_t, the limits of the sample ACF are in general random. This is in contrast to the asymptotic theory for the sample autocorrelations of a linear process as treated in [11] and [12] (see [4], Section 13.3). In this case, the sample ACF estimates the "model ACF" with scaling that is larger than $\sqrt{n}$. Even though the model ACF does not exist in the infinite variance case, it can be defined, at least for linear processes, through the filter coefficients. Specifically, the model ACF is defined as the ACF of a Gaussian linear process with the same coefficients as (X_t).

The phenomenon of random limits of the ACF was observed earlier in the context of infinite variance bilinear processes (see [13] and [30]), and our Theorem 3.5 confirms that deterministic limits are more the exception than the rule for the sample ACF of a nonlinear stationary sequence. Theorem 3.5 also treats the finite variance but infinite fourth moment case. In this situation and providing the process is ergodic, the sample ACF estimates the ACF consistently, but the limit results show that the asymptotic rate is slower than $\sqrt{n}$. (For a linear process with finite variance and infinite fourth moment, the sample ACF is still asymptotically normal with scaling $n^{1/2}$.) Therefore the values of the sample ACF of an infinite fourth moment stationary process have to be treated with enormous care; the confidence bands in this case are wider than the classical plus or minus $2\sigma/\sqrt{n}$-bounds. This is, on the one hand, caused by the infinite variance limit distribution and, on the other hand, by the slow rate of convergence, which is slower the closer one comes to an infinite second moment.

Bearing these facts in mind, we apply the general theory to the simple model of a stationary ARCH(1) for illustrative purposes. This process is

known to have regularly varying marginals (see [23] and [18]). The importance of ARCH processes for financial modelling has been mentioned before and will be explained in more detail in Section 4 where we also give the general theory for the sample ACVF and ACF of an ARCH(1) process, its absolute values and squares.

The principal objective of this paper is to develop general methodology for deriving distributional convergence of the sample ACF and ACVF for heavy-tailed stationary sequences. We restrict ourselves to qualitative results. In future work we intend to identify the limit distributions, thus making the results directly applicable for estimation and testing purposes. From the method of proof it will become more transparent that the proposed technique allows for very general classes of stationary processes. This includes linear and multilinear processes, as well as multivariate processes, which can be written as the solution to certain stochastic recurrence equations. This will be the content of a forthcoming paper.

It is worth remarking that the asymptotic behavior of the sample ACF for some *linear* processes, such as multivariate and periodic moving averages, may be similar to that obtained for nonlinear processes. For example, in a multivariate moving average process driven by noise that is regularly varying, the sample cross-correlation function has a random limit if the process has an infinite variance [10]. On the other hand, if the variance is finite but the fourth moment is infinite, then the asymptotic scaling of the sample cross-correlation function has the same form as specified in the Theorem 3.5 [9]. Analogous behavior for the sample ACF of periodic moving averages was established in [1].

The paper is organized as follows. In Section 2 we sketch the necessary point process theory. In most parts it is analogous to the corresponding results in [8]; we only indicate the changes required for the results to go through in the multivariate setting. In Section 3 we give the limit theory for the sample ACVF and ACF of a stationary sequence. Section 4 is entirely devoted to the ARCH model, where the theory of the previous sections is demonstrated. In the Appendix we prove some useful facts (e.g., strong mixing and joint regular variation) about the ARCH(1) process.

2. Some point process theory. We consider a strictly stationary sequence $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ of random row vectors with values in $\mathbb{R}^m$. For simplicity, we write $\mathbf{X} = \mathbf{X}_0 = (X_1, \ldots, X_m)$.

The regular variation condition. Assume that the distribution of X is *jointly regularly varying with index* $\alpha > 0$. This means that there exists a sequence of constants x_n and a random vector $\boldsymbol{\theta} \in \mathbb{S}^{m-1}$ a.s., where $\mathbb{S}^{m-1}$ denotes the unit sphere in $\mathbb{R}^m$ with respect to the norm $|\cdot|$, such that

$$(2.1) \qquad nP(|\mathbf{X}| > tx_n, \mathbf{X}/|\mathbf{X}| \in \cdot) \to_v t^{-\alpha} P(\boldsymbol{\theta} \in \cdot), \qquad t > 0$$

where $\to_v$ denotes vague convergence on $\mathbb{S}^{m-1}$.

This is the same as

$$\frac{P(|\mathbf{X}| > tx, \mathbf{X}/|\mathbf{X}| \in \cdot)}{P(|\mathbf{X}| > x)} \to_v t^{-\alpha} P(\boldsymbol{\theta} \in \cdot), \qquad t > 0$$

as $x \to \infty$ (cf. [32] and [14]). For our application, it is natural to take $|\cdot|$ to be the max-norm in $\mathbb{R}^m$, that is,

$$|\mathbf{x}| = |(x_1, \ldots, x_m)| = \max_{i=1,\ldots,m} |x_i|.$$

Note that regular variation of $\mathbf{X}$ in $\mathbb{R}^m$ implies regular variation of $|\mathbf{X}|$. For further information on multivariate regular variation we refer to [2]. Vague convergence of measures is treated in detail in [22].

Preliminaries on point processes. We follow the point process theory in [22]. The state space of the point processes considered is $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. Write $\mathscr{B}$ for the collection of bounded Borel sets in $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$. ("Bounded" here means bounded away from the origin.) Let $\mathscr{F}$ be the collection of bounded nonnegative continuous functions on $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$ with bounded support and let $\mathscr{F}_s$ be the collection of bounded nonnegative step functions on $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$ with bounded support.

Write $\mathscr{M}$ for the collection of Radon counting measures on $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$ with null measure o. This means $\mu \in \mathscr{M} \setminus \{o\}$ if and only if μ is of the form $\sum_{i=1}^{\infty} n_i \varepsilon_{\mathbf{x}_i}$, where $n_i \in \{1, 2, \ldots\}$, the points $\mathbf{x}_i$ are distinct and $\bigvee_{i=1}^{\infty} |\mathbf{x}_i| < \infty$. Let $\mathscr{M}_y \subset \mathscr{M}$ be the collection of measures μ such that $\mu(\{\mathbf{x}: |\mathbf{x}| > y\}) > 0$. Note that $\mathscr{M}_0 = \mathscr{M} \setminus \{o\}$.

The mixing condition. Let (a_n) be a sequence of positive numbers such that

(2.2) $$nP(|\mathbf{X}| > a_n) \to 1, \qquad n \to \infty.$$

In particular, one can choose a_n as the $(1 - n^{-1})$-quantile of $|\mathbf{X}|$. Since $|\mathbf{X}|$ is regularly varying, $a_n = n^{1/\alpha} L(n)$ for some slowly varying function $L(x)$.

We say that the *condition $\mathscr{A}(a_n)$* holds for $(\mathbf{X}_t)$ if there exists a sequence of positive integers (r_n) such that $r_n \to \infty$, $k_n = [n/r_n] \to \infty$ as $n \to \infty$ and

(2.3) $$E \exp\left\{ - \sum_{t=1}^{n} f(\mathbf{X}_t/a_n) \right\} - \left(E \exp\left\{ - \sum_{t=1}^{r_n} f(\mathbf{X}_t/a_n) \right\} \right)^{k_n} \to 0,$$

$$n \to \infty, \forall f \in \mathscr{F}_s.$$

The convergence in (2.3) is not required to be uniform in f. This is indeed a very weak condition and is implied by many known mixing conditions, in particular the strong mixing condition. The condition $\mathscr{A}(a_n)$ is similar in spirit to condition Δ used in extreme value theory; see, for example, [25]. However, these conditions are not directly comparable since Δ is defined in terms of probabilities of events in restricted classes of σ-fields, whereas $\mathscr{A}(a_n)$ is specified in terms of Laplace transforms of point processes. Lemma

2.4.2 in [25] shows that Δ and $\mathscr{A}(a_n)$ are close indeed. Condition $\mathscr{A}(a_n)$ is independent of the particular choice of (a_n): if both (a_n) and (a'_n) obey (2.2), then $\mathscr{A}(a_n)$ holds if and only if $\mathscr{A}(a'_n)$ does. The mixing condition $\mathscr{A}(a_n)$ was introduced in [8] for the case of real-valued stationary sequences.

A corresponding statement can be made for the weak convergence of the sequence of point processes

$$N_n = \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t / a_n}, \qquad n = 1, 2, \dots.$$

Define

$$\tilde{N}_n = \sum_{i=1}^{k_n} \tilde{N}_{r_n, i},$$

with $\tilde{N}_{r_n, i}$, $i = 1, \dots, k_n$, iid distributed as $\tilde{N}_{r_n, 0} = \sum_{t=1}^{r_n} \varepsilon_{\mathbf{X}_t / a_n}$. By virtue of (2.3), a Laplace transform argument shows that condition $\mathscr{A}(a_n)$ implies that (N_n) converges weakly if and only if $(\tilde{N}_n)$ does, and they have the same limit. This is an important ingredient in the proof of Lemma 2.1.

Main results on point process convergence. In what follows, we give a series of results which were proved in [8] in the particular case $m = 1$. We formulate them here since they are the theoretical basis for the following sections; their proofs are analogous to the one-dimensional case and are omitted except when essential modifications due to the multivariate nature of the process are required. We commence with the analogues to Lemmas 2.1 and 2.2 in [8].

LEMMA 2.1. *Suppose that* $(\mathbf{X}_t)$ *obeys conditions* $\mathscr{A}(a_n)$ *and* (2.1). *The relation* $N_n \rightarrow_d N \neq o$ *holds if and only if there exists a nonnull measure* λ *on* $\mathscr{M}_0$ *with* $\int (1 - \exp\{-\mu(B)\})\lambda(d\mu) < \infty$ *for all* $B \in \mathscr{B}$, *such that*

$$\int (1 - e^{-\mu f}) \lambda_n(d\mu) \rightarrow \int (1 - e^{-\mu f}) \lambda(d\mu), \qquad f \in \mathscr{F},$$

where $\lambda_n = k_n(P \circ \tilde{N}_{r_n, 0}^{-1})$ *and* $\mu f = \int f d\mu$. *The point process* N *is infinitely divisible with Laplace transform* $\exp\{-\int (1 - \exp\{-\mu f\})\lambda(d\mu)\}$, N *has no fixed atoms and the support of* $P \circ N^{-1}$ *and that of* λ *are both contained in* $\mathscr{M}_0$. *Moreover, the canonical measure* λ *of* N *has the scaling property* $\lambda(\cdot) = \sigma^\alpha \lambda(\sigma^{-1} \cdot)$, $\sigma > 0$.

Define

$$(2.4) \quad \tilde{\mathscr{M}} = \left\{ \mu \in \mathscr{M} : \mu(\{\mathbf{x} : |\mathbf{x}| > 1\}) = 0 \quad \text{and} \quad \mu(\{\mathbf{x} : \mathbf{x} \in \mathbb{S}^{m-1}\}) > 0 \right\},$$

and let $\mathscr{B}(\tilde{\mathscr{M}})$ be the Borel σ-field of $\tilde{\mathscr{M}}$. For $\mu = \sum_{i=1}^{\infty} n_i \varepsilon_{\mathbf{x}_i} \in \mathscr{M}_0$, let $x_{\mu} = \max_i |\mathbf{x}_i|$. Define a mapping on $\mathscr{M}_0$ by $T_1: \mu \to (x_{\mu}, \mu(x_{\mu} \cdot))$, that is,

$$(2.5) \qquad T_1\left(\sum_{i=1}^{\infty} n_i \varepsilon_{\mathbf{x}_i} \right) = \left(\max_i |\mathbf{x}_i|, \sum_{i=1}^{\infty} n_i \varepsilon_{\mathbf{x}_i / \max_j |\mathbf{x}_j|} \right).$$

It is a bicontinuous bijection with range $\mathbb{R}_+ \times \tilde{\mathscr{M}}$, where $\mathbb{R}_+ = (0, \infty)$.

The following result is the multivariate analogue of Theorem 2.3 in [8].

THEOREM 2.2.   *Assume that* $(\mathbf{X}_t)$ *obeys the mixing condition* $\mathscr{A}(a_n)$ *and the regular variation condition* (2.1). *If* $N_n \to_d N \neq o$, *then* N *is infinitely divisible with canonical measure* λ *on* $\mathscr{M}_0$ *and* $T_1(N_n) \to_d T_1(N)$. *The limiting process* $T_1(N)$ *has canonical measure* $\lambda \circ T_1^{-1} = \nu \times Q$, *where* Q *is a probability measure on* $(\tilde{\mathscr{M}}, \mathscr{B}(\tilde{\mathscr{M}}))$,

$$\nu(dy) = \gamma \alpha y^{-\alpha-1} I_{\mathbb{R}_+}(y)\, dy \quad \text{and}$$

$$\gamma = \lambda(\{ \mu: \mu(\{\mathbf{x}: |\mathbf{x}| > 1\}) > 0\}) \in (0, 1].$$

In this case, the Laplace transform of N *is*

$$(2.6) \quad \exp\left\{ -\int_0^{\infty} \int_{\tilde{\mathscr{M}}} \left(1 - \exp\left(-\int f(yz) \mu(dz) \right) \right) Q(d\mu) \nu(dy) \right\}, \qquad f \in \mathscr{F}.$$

REMARK 2.3.   Notice that

$$\gamma = \lambda(\{ \mu: \mu(\{\mathbf{x}: |\mathbf{x}| > 1\}) > 0\}) = -\ln P(N(\{\mathbf{x}: |\mathbf{x}| > 1\}) = 0)$$

is the *extremal index* of $(|\mathbf{X}_t|)$; see [24].

The following result is the analogue to Corollary 2.4 in [8]. It follows from a comparison of the Laplace transforms of the two point processes involved.

COROLLARY 2.4 (Cluster representation).   *Let* N *be the limiting point process in Theorem* 2.2. *Then* N *is identical in law to the point process*

$$(2.7) \qquad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}},$$

where $\sum_{i=1}^{\infty} \varepsilon_{P_i}$ *is a Poisson process on* $\mathbb{R}_+$ *with intensity measure* ν, *independent of the sequence of iid point processes* $\sum_{j=1}^{\infty} \varepsilon_{\mathbf{Q}_{ij}}$, $i \geq 1$, *with joint distribution* Q *on* $(\tilde{\mathscr{M}}, \mathscr{B}(\tilde{\mathscr{M}}))$.

REMARK 2.5.   Representation (2.7) shows nicely the structure of the points of the limiting point process: P_i describes the radial part of the points and $\mathbf{Q}_{ij}$ the spherical part. Their independence is a consequence of the factorization of the limiting measure in the multivariate regular variation condition (2.1).

Now we provide a useful sufficient condition for the weak convergence of the point processes N_n toward N. It is needed later in the proof of Theorem 2.8.

THEOREM 2.6. *Assume that* $(\mathbf{X}_t)$ *obeys* $\mathscr{A}(a_n)$ *and* (2.1), *and the following two conditions hold for every* $y > 0$:

$$(2.8) \qquad k_n P\left(\bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n y\right) \to \gamma y^{-\alpha} = \nu(y, \infty),$$

$$(2.9) \qquad P\left(\sum_{t=1}^{r_n} \varepsilon_{\mathbf{X}_t / \vee_{j=1}^{r_n} |\mathbf{X}_j|} \in \cdot \,\middle|\, \bigvee_{t=1}^{r_n} |\mathbf{X}_t| > y a_n\right) \to_w Q,$$

where $\to_w$ *denotes weak convergence and* γ *is given in Theorem 2.2. Then* $N_n \to_d N$.

REMARK 2.7. In the case $m = 1$, conditions (2.8) and (2.9) are known to be necessary and sufficient for $N_n \to_d N$; see [8], Theorem 2.5.

PROOF. Recall the definition of the measures λ_n from Lemma 2.1 and of $\mathscr{M}_y$ from the beginning of this section. For $y > 0$ fixed, define the (conditional) probability measures $P_{n,y}$ on $\mathscr{M}_0$,

$$P_{n,y}(\cdot) = \frac{\lambda_n(\cdot \cap \mathscr{M}_y)}{\lambda_n(\mathscr{M}_y)}$$

$$= \frac{P\left(\sum_{t=1}^{r_n} \varepsilon_{\mathbf{X}_t / a_n} \in \cdot, \bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n y\right)}{P\left(\bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n y\right)}$$

$$= P\left(\sum_{t=1}^{r_n} \varepsilon_{\mathbf{X}_t / a_n} \in \cdot \,\middle|\, \bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n y\right).$$

Fix $f \in \mathscr{F}$ and suppose that its support is contained in $\{\mathbf{x}: |\mathbf{x}| > y\}$ for some $y = y_f > 0$. Recall the definition of the mapping T_1 from (2.5). Then for any $z > 0$,

$$\left(P_{n,y} \circ T_1^{-1}\right)((z, \infty) \times \cdot) = \frac{P\left(\sum_{t=1}^{r_n} \varepsilon_{\mathbf{X}_t / \vee_{j=1}^{r_n} |\mathbf{X}_j|} \in \cdot, \bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n(y \vee z)\right)}{P\left(\bigvee_{t=1}^{r_n} |\mathbf{X}_t| > a_n y\right)}.$$

By (2.8) and (2.9), the right-hand side converges weakly to

$$Q(\cdot) \nu((y \vee z), \infty) / \nu(y, \infty).$$

It follows that

$$P_{n,y} \circ T_1^{-1} \to_w \frac{\nu(\cdot \cap (y, \infty)) \times Q}{\nu(y, \infty)} \quad \text{on } \mathbb{R}_+ \times \tilde{\mathscr{M}},$$

and since T_1^{-1} is continuous, we have

$$P_{n,y}(\cdot) \to_w \frac{((\nu \times Q) \circ T_1)(\cdot \cap \mathscr{M}_y)}{\nu(y, \infty)}.$$

Defining $\lambda = (\nu \times Q) \circ T_1$, the latter relation can be rewritten as

$$P_{n,y}(\cdot) \rightarrow_w \frac{\lambda(\cdot \cap \mathcal{M}_y)}{\lambda(\mathcal{M}_y)} =: P_y(\cdot).$$

Since $f \in \mathcal{F}$,

$$\int_{\mathcal{M}_y} e^{-\mu f} P_{n,y}(d\mu) \rightarrow \int_{\mathcal{M}_y} e^{-\mu f} P_y(d\mu).$$

Since the support of f is contained in $\{\mathbf{x} : |\mathbf{x}| > y\}$, we have

$$\int (1 - e^{-\mu f}) \lambda_n(d\mu) \rightarrow \int (1 - e^{-\mu f}) \lambda(d\mu),$$

and the result now follows from Lemma 2.1. □

The following result is the analogue to Theorem 2.7 in [8].

THEOREM 2.8. *Assume that* $(\mathbf{X}_t)$ *is a stationary sequence of random vectors for which all finite-dimensional distributions are jointly regularly varying with index* $\alpha > 0$. *To be specific, let* $(\boldsymbol{\theta}_{-k}, \ldots, \boldsymbol{\theta}_k)$ *be the* $(2k+1)m$-*dimensional random row vector with values in the unit sphere* $\mathbb{S}^{(2k+1)m-1}$ *that appears in the definition of joint regular variation of* $(\mathbf{X}_{-k}, \ldots, \mathbf{X}_k)$, $k \geq 0$. [*Note: the defining property* (2.1) *has to be applied to this* $(2k+1)m$-*dimensional setting. In particular,* $|\cdot|$ *has to be interpreted as the* max-norm *in* $\mathbb{R}^{(2k+1)m}$]. *Assume that condition* $\mathscr{A}(a_n)$ *holds for* $(\mathbf{X}_t)$ *and that*

(2.10) $$\lim_{k \to \infty} \limsup_{n \to \infty} P\left(\bigvee_{k \leq |t| \leq r_n} |\mathbf{X}_t| > a_n y \,\middle|\, |\mathbf{X}_0| > a_n y \right) = 0, \qquad y > 0.$$

Then the limit

(2.11) $$\gamma = \lim_{k \to \infty} E\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^k |\boldsymbol{\theta}_j^{(k)}|^\alpha \right)_+ \Big/ E|\boldsymbol{\theta}_0^{(k)}|^\alpha$$

exists. If $\gamma = 0$, *then* $N_n \rightarrow_d o$; *if* $\gamma > 0$, *then* $N_n \rightarrow_d N \neq o$, *where, using the representation* $\lambda \circ T^{-1} = \nu \times Q$ *described in Theorem* 2.2, $\nu(dy) = \gamma \alpha \, y^{-\alpha-1} \, dy$, *and* Q *is the weak limit of*

$$E\left[\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^k |\boldsymbol{\theta}_j^{(k)}|^\alpha \right)_+ I. \left(\sum_{|t| \leq k} \varepsilon_{\boldsymbol{\theta}_t^{(k)}} \right) \right] \Big/ E\left(|\boldsymbol{\theta}_0^{(k)}|^\alpha - \bigvee_{j=1}^k |\boldsymbol{\theta}_j^{(k)}|^\alpha \right)_+$$

as $k \to \infty$ *which exists.*

For the proof of Theorem 2.8, the analogues of Lemmas 2.8 and 2.9 in [8] have to be modified for the multivariate setting. This is a consequence of the fact that these results only make use of the absolute values of the quantities $\mathbf{X}_t$. Applying these two lemmas, one can follow the lines of the proof of Theorem 2.7 in [8] to derive Theorem 2.8. The crucial difference to the

one-dimensional case as treated in [8] is the definition of the sets A in Lemma 2.8:

$$A = \left\{ \mu \in \tilde{\mathcal{M}} : \mu(\{\mathbf{x} : b_i < |\mathbf{x}| \leq 1, \mathbf{x}/|\mathbf{x}| \in B_i\}) \geq t_i, i = 1, \ldots, l \right\},$$

for some $l \geq 1$, $t_i \geq 1$, $0 < b_i < 1$, Borel sets $B_i \in \mathbb{S}^{m-1}$, $i = 1, \ldots, l$. In the multivariate setting, the collection of sets A determines the measure Q on $\tilde{\mathcal{M}}$.

REMARK 2.9. In the above theory, $|\cdot|$ was the max-norm in finite-dimensional Euclidean space. It is the natural norm to consider for our applications. However, the choice of the norm is not essential for the above theory. Indeed, the max-norm $|\cdot|$ can be replaced by any other norm in the regular variation and mixing conditions. Then Theorems 2.2 and 2.6 remain valid without any changes, and in Theorem 2.8 the regular variation condition on $\mathbf{Y}_k = (\mathbf{X}_{-k}, \ldots, \mathbf{X}_k)$, $k \geq 0$, has to be in terms of the norm in $\mathbb{R}^{(2k+1)m}$ which is consistent with the norm used in the definition (2.1) of regular variation for the marginal distribution $\mathbf{X}$.

The point process theory of this section can also be extended to an infinite-dimensional setting. For example, if $(\mathbf{X}_t)$ is a stationary sequence of random elements assuming values in a separable locally compact Banach space $(\mathbf{B}, |\cdot|)$, then the point process theory in [22] is still applicable. In this case, $\mathbf{B} \setminus \{\mathbf{0}\}$ is the state space and is a subset of a Polish space. The mixing and regular variation conditions on $(\mathbf{X}_t)$ remain the same if one interprets $|\cdot|$ as the norm in $\mathbf{B}$, and $\mathbb{S}^{m-1}$ has to be replaced by the unit sphere $\mathbb{S}_{\mathbf{B}}$ in $\mathbf{B}$. Again, Theorems 2.2 and 2.6 remain valid without any changes. However, the definition of regular variation of $\mathbf{Y}_k$ in Theorem 2.8 needs to be modified: interpret $\mathbf{Y}_k$ as an element of $\mathbf{B}^{2k+1}$ endowed with the norm $\|\mathbf{x}\| = \vee_{i=-k}^{k} |\mathbf{x}_i|$. With this norm, one can define regular variation in $\mathbf{B}^{2k+1}$ in the same way as in (2.1), and all the steps in the proof of Theorem 2.8 remain valid without any changes. As a consequence, a great part of the results in [8], Sections 3 and 4, on partial sum convergence and large deviation probabilities remains valid for stationary $\mathbf{B}$-valued sequences $(\mathbf{X}_t)$ with marginal distribution in the domain of attraction of an α-stable law in $\mathbf{B}$, $\alpha < 2$. Clearly, every step in the proof of those results has to be checked carefully; for example, $|\cdot|$ has to be interpreted as the norm in $\mathbf{B}$, characteristic functions are characteristic functionals, expectations have to be interpreted in the Bochner sense, care has to be taken when moment inequalities are applied, and so on. We omit further details.

3. Limit theory for the sample ACF of a stationary process. We start this section with an elementary but powerful result on point process convergence, where the points of the processes are products of random variables. It will turn out to be the basis for the results on weak convergence of sample autocovariances for various stationary processes.

PROPOSITION 3.1.   *Let* (X_t) *be a strictly stationary sequence such that* $(\mathbf{X}_t) = ((X_t, \ldots, X_{t+m}))$ *satisfies* (2.1) *for some* $m \geq 0$ *and*

$$N_n = \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t / a_n} \to_d N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}},$$

where the points $\mathbf{Q}_{ij} = (Q_{ij}^{(0)}, \ldots, Q_{ij}^{(m)})$ *and* P_i *are defined in Corollary* 2.4 *(with* $m + 1$*). Then*

(3.1)

$$\hat{\mathbf{N}}_n = \left(\hat{N}_{n,h} \right)_{h=0,\ldots,m} = \left(\sum_{t=1}^{n} \varepsilon_{a_n^{-2} X_t X_{t+h}} \right)_{h=0,\ldots,m}$$

$$\to_d \hat{\mathbf{N}} = \left(\hat{N}_h \right)_{h=0,\ldots,m} = \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}} \right)_{h=0,\ldots,m}.$$

The components of $\hat{\mathbf{N}}_n$ *and* $\hat{\mathbf{N}}$ *are considered as point processes on* $\overline{\mathbb{R}} \setminus \{0\}$. *This means that points are not included in the point processes if* $X_t X_{t+h} = 0$ *or* $Q_{ij}^{(0)} Q_{ij}^{(h)} = 0$.

REMARK 3.2.   From the proof below it is immediate that Proposition 3.1 can be extended to point processes of cross products of two different stationary sequences (X_t) and (Y_t), say. For example, if the sequence of random vectors $(X_t, \ldots, X_{t+m}, Y_t, \ldots, Y_{t+m})$ satisfies $\mathscr{A}(a_n)$ and (2.1), then the vector of point processes

$$\left(\sum_{t=1}^{n} \varepsilon_{a_n^{-2} X_t Y_{t+h}} \right)_{h=0,\ldots,m}$$

converges in distribution, and the structure of the limit is similar to the one in (3.1). This limit result can be exploited to obtain joint convergence of the sample cross-covariances and sample cross-correlations of the X- and Y-processes. One can follow the pattern of proof given below in the case of sample autocovariances and autocorrelations of the process (X_t). In the same way, one can also derive the limit theory for the sample cross-covariances and cross-correlations for an arbitrary number of stationary sequences.

PROOF.   We show marginal convergence of $\hat{N}_{n,h}$ to $\hat{N}_h$, the joint convergence being a straightforward extension of the argument. For $h = 0, \ldots, m$, define

$$\hat{T}_h : \mathbf{x} = (x_0, \ldots, x_m) \in \overline{\mathbb{R}}^{m+1} \setminus \{\mathbf{0}\} \to x_0 x_h.$$

Then $\hat{N}_{n,h} = N_n \circ \hat{T}_h^{-1}$ and $\hat{N}_h = N \circ \hat{T}_h^{-1}$ with the null points excluded from $\hat{N}_{n,h}$ and $\hat{N}_h$, as mentioned in the formulation of the proposition. Let $I_1, \ldots, I_k$ be bounded intervals in $\overline{\mathbb{R}} \setminus \{0\}$ (i.e., bounded away from the origin). As such, they are elements of a DC-ring in the sense of [22], page 11. By Theorem 4.2 in the same reference, it suffices to show convergence of the finite-dimensional distributions, that is,

(3.2) $\left(\hat{N}_{n,h}(I_1), \ldots, \hat{N}_{n,h}(I_k) \right) \to_d \left(\hat{N}_h(I_1), \ldots, \hat{N}_h(I_k) \right).$

However, the sets $\hat{T}_h^{-1}(I_j)$ are also bounded away from zero in $\overline{\mathbb{R}}^{m+1} \setminus \{\mathbf{0}\}$, and they are N-continuity sets, that is, $N(\partial \hat{T}_h^{-1}(I_j)) = 0$ a.s. This follows from the fact that N does not have fixed atoms; compare Lemma 2.1 in combination with Theorem 2.2 and Corollary 2.4. Now (3.2) follows from Lemma 4.4 in [22]. This completes the proof. $\square$

The convergence of the sample autocorrelations can be derived from the following result, which is a generalization of Theorem 3.1 in [8] to the multivariate case.

PROPOSITION 3.3. *Let* $(\mathbf{Z}_t)$ *be a strictly stationary sequence of random vectors in* $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$ *and* (b_n) *be a sequence of positive numbers such that*

$$\left(\sum_{t=1}^{n} \varepsilon_{Z_t^{(h)}/b_n} \right)_{h=1,\ldots,m} \to_d \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}^{(h)}} \right)_{h=1,\ldots,m},$$

where (P_i) *are the points of a Poisson process on* $\mathbb{R}_+$, *independent of the iid point processes* $\sum_{j=1}^{\infty} \varepsilon_{\mathbf{Q}_{ij}}$, $i \geq 1$, *on* $\overline{\mathbb{R}}^m \setminus \{\mathbf{0}\}$. *Also assume that* $nP(\mathbf{Z}/b_n \in \cdot) \to_v \tau(\cdot)$ *for some measure* τ *on* $\mathbb{R}^m \setminus \{\mathbf{0}\}$ *which is the Lévy measure of an* α-*stable random vector,* $\alpha \in (0, 2)$, *with values in* $\mathbb{R}^m$. *Let*

$$\mathbf{S}_0 = \mathbf{0}, \qquad \mathbf{S}_n = \mathbf{Z}_1 + \cdots + \mathbf{Z}_n, \qquad n \geq 1,$$

and for any Borel set B in $\mathbb{R}$, set

$$\mathbf{S}_n B = \left(S_n^{(h)}(B) \right)_{h=1,\ldots,m},$$

where

$$S_n^{(h)}(B) = b_n^{-1} \sum_{t=1}^{n} Z_t^{(h)} I_B \left(|Z_t^{(h)}|/b_n \right), \qquad n \geq 1.$$

(i) *If* $\alpha \in (0, 1)$, *then*

$$b_n^{-1} \mathbf{S}_n \to_d \mathbf{S} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i \mathbf{Q}_{ij},$$

and $\mathbf{S}$ *has an* α-*stable distribution.*
    (ii) *If* $\alpha \in [1, 2)$ *and for all* $\delta > 0$,

(3.3) $$\lim_{y \to 0} \limsup_{n \to \infty} P\left(|S_n^{(h)}(0, y] - ES_n^{(h)}(0, y]| > b_n \delta \right) = 0,$$

$h = 1, \ldots, m$, *then*

$$b_n^{-1}\left(\mathbf{S}_n - E\mathbf{S}_n(0, 1] \right) \to_d \mathbf{S},$$

where $\mathbf{S}$ *is the distributional limit of*

$$\left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i Q_{ij}^{(h)} I_{(y, \infty]} \left(|P_i Q_{ij}^{(h)}| \right) - \int_{\mathbf{x}:\, y < |x_h| \leq 1} x_h \tau(d\mathbf{x}) \right)_{h=1,\ldots,m}$$

as $y \to 0$, *which exists and has an* α-*stable distribution.*

REMARK 3.4. As in the univariate case, $\mathbf{S}_n$ can be centered by $E\mathbf{S}_n$ provided that $\alpha > 1$ (see Remark 3.1 in [8]). In this case, it follows from the vague convergence that $b_n^{-1}(E\mathbf{S}_n - E\mathbf{S}_n(0, 1]) \to (\int_{\mathbf{x}:\, 1 < |x_h|} x_h \tau(d\mathbf{x}))_{h=1,\ldots,m}$.

PROOF. The proposition is the multivariate analogue to Theorem 3.1 in [8]. If $\alpha \in (0, 1)$, one can follow the lines of its proof. If $\alpha \in [1, 2)$, a multivariate characteristic function argument replaces the argument on pages 897 and 898 to show that $(\mathbf{S}_n - E\mathbf{S}_n(0, 1])$ converges in distribution. To characterize the limit as α-stable, it suffices to show that every linear combination is α-stable (cf. [31], Theorem 2.1.5(c)). This follows by the argument on page 898 in [8]. □

Combining Propositions 3.1 and 3.3, we can derive the asymptotic limit behavior of the sample autocovariances and autocorrelations of a stationary sequence (X_t). Construct from this process the strictly stationary m-dimensional processes $(\mathbf{X}_t(m)) = ((X_t, \ldots, X_{t+m}))$, $m \geq 0$. Define the *sample autocovariance function*

$$\gamma_{n,X}(h) = n^{-1} \sum_{t=1}^{n-h} X_t X_{t+h}, \qquad h \geq 0,$$

and the corresponding *sample autocorrelation function*

$$\rho_{n,X}(h) = \gamma_{n,X}(h)/\gamma_{n,X}(0), \qquad h \geq 1.$$

We also write

$$\gamma_X(h) = EX_0 X_h \quad \text{and} \quad \rho(h) = \gamma_X(h)/\gamma_X(0), \qquad h \geq 0,$$

for the autocovariances and autocorrelations, respectively, of the sequence (X_t) if these quantities exist. Mean-corrected versions of both the sample and model ACVF can also be considered (see Remark 3.6).

THEOREM 3.5. *Assume that (X_t) is a strictly stationary sequence of random variables and that for some fixed $m \geq 0$, $(\mathbf{X}_t(m))$ satisfies the regular variation condition (2.1) (with $m + 1$) and $N_n = \sum_{t=1}^n \varepsilon_{\mathbf{X}_t/a_n} \to_d N = \sum_{i=1}^\infty \sum_{j=1}^\infty \varepsilon_{P_i \mathbf{Q}_{ij}}$, where the points $\mathbf{Q}_{ij} = (Q_{ij}^{(0)}, \ldots, Q_{ij}^{(m)})$ and P_i are as defined in Corollary 2.4.*

(i) *If $\alpha \in (0, 2)$, then*

$$\left(na_n^{-2}\gamma_{n,X}(h)\right)_{h=0,\ldots,m} \to_d (V_h)_{h=0,\ldots,m},$$

$$\left(\rho_{n,X}(h)\right)_{h=1,\ldots,m} \to_d (V_h/V_0)_{h=1,\ldots,m},$$

where

$$V_h = \sum_{i=1}^\infty \sum_{j=1}^\infty P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)}, \qquad h = 0, \ldots, m.$$

The vector $(V_0, \ldots, V_m)$ *is jointly* $\alpha/2$-*stable in* $\mathbb{R}^{m+1}$.

(ii) *If* $\alpha \in (2, 4)$ *and for* $h = 0, \ldots, m$,

$$
(3.4) \qquad \lim_{\epsilon \to 0} \limsup_{n \to \infty} \mathrm{var}\left(a_n^{-2} \sum_{t=1}^{n-h} X_t X_{t+h} I_{\{|X_t X_{t+h}| \le a_n^2 \epsilon\}} \right) = 0,
$$

then

$$
(3.5) \qquad \left(n a_n^{-2} (\gamma_{n,X}(h) - \gamma_X(h)) \right)_{h=0,\ldots,m} \to_d (V_h)_{h=0,\ldots,m},
$$

where $(V_0, \ldots, V_m)$ *is the distributional limit of*

$$
\left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^{(0)} Q_{ij}^{(h)} I_{(\epsilon, \infty]}\left(P_i^2 |Q_{ij}^{(0)} Q_{ij}^{(h)}| \right) - \int_{B_{\epsilon, h}} x_0 x_h \tau(d\mathbf{x}) \right)_{h=0,\ldots,m},
$$

where

$$
B_{\epsilon, h} = \left\{ \mathbf{x} \in \mathbb{R}^{m+1} : \epsilon < |x_0 x_h| \right\},
$$

and τ *is the measure as defined in Proposition* 3.3 *for* $\mathbf{Z}_t = (X_t, \ldots, X_{t+m})$. *Moreover*,

$$
(3.6) \quad \left(n a_n^{-2} (\rho_{n,X}(h) - \rho_X(h)) \right)_{h=1,\ldots,m} \to_d \gamma_X^{-1}(0)(V_h - \rho_X(h)V_0)_{h=1,\ldots,m}.
$$

PROOF. Part (i) is immediate from Propositions 3.1 and 3.3(i) and the continuous mapping theorem. The convergence of $\gamma_{n,X}$ in part (ii) is a direct consequence of Proposition 3.3(ii) and Remark 3.4. To prove (3.6), note that from (3.5),

$$
\gamma_{n,X}(h) = \gamma_X(h) + O_P(n^{-1} a_n^2).
$$

Since $n^{-1} a_n^2 \to 0$, we obtain, after applying a Taylor series expansion to the function $f(x, y) = x/y$,

$$
\rho_{n,X}(h) = \rho_X(h) + (\gamma_{n,X}(h) - \gamma_X(h))/\gamma_X(0)
$$
$$
- (\gamma_{n,X}(0) - \gamma_X(0))\gamma_X(h)/\gamma_X^2(0) + o_P(n^{-1} a_n^2).
$$

The conclusion (3.6) is now immediate from (3.5), the continuous mapping theorem and an application of Slutsky's theorem. □

REMARK 3.6. The limit distributions of both $\gamma_{n,X}(h)$ and $\rho_{n,X}(h)$ in (3.5) and (3.6) of Theorem 3.5 are $\alpha/2$-stable. The case $\alpha = 2$ can also be included but it leads to some tedious considerations of the centering constants and is therefore omitted. One can also consider the mean-corrected versions of the sample ACVF and ACF; the same arguments as above show that the limit theory does not change.

REMARK 3.7. If $\alpha \in (0, 2)$, the sample autocorrelations have a constant limit if and only if $V_h = \rho_h V_0$ for some constants ρ_h. The quantities ρ_h play a role similar to the autocorrelations, as a measure of dependence in the sequence $(\mathbf{X}_t)$. If $\alpha \in (2, 4)$, the sample autocorrelations $\rho_{n,X}(h)$ are clearly

consistent estimates of the autocorrelations $\rho_X(h)$, but with rate $na_n^{-2} = n^{1-2/\alpha}L(n)$ for some slowly varying L. This implies that the asymptotic confidence bands are wider than in the case $EX^4 < \infty$, where $n^{1/2}$ is the rate implied by the CLT. An improvement of this rate is possible only if $V_h = \rho_X(h)V_0$.

4. The sample ACVF and ACF of an ARCH(1) process. In this section we consider the ARCH(1) process defined through the recursions,

$$(4.1) \qquad X_t = \left(\beta + \lambda X_{t-1}^2 \right)^{1/2} Z_t, \qquad t \in \mathbb{Z}, \, (Z_t) \text{ iid } N(0,1),$$

where β and λ are positive parameters. We follow the theory in [17], Section 8.4. If $\lambda \in (0, 2e^E)$, where $E = 0.5772\ldots$ is Euler's constant, then $2e^E = 3.5620\ldots$, and there exists a stationary solution (X_t) to (4.1) with marginal distribution satisfying

$$(4.2) \qquad X_0^2 \stackrel{d}{=} \frac{\beta}{\lambda} \sum_{t=1}^{\infty} \prod_{j=1}^{t} \left(\lambda Z_j^2 \right).$$

It follows from work by Kesten [23] and Goldie [18] that

$$(4.3) \qquad P(X > x) \sim c_\alpha x^{-\alpha}, \qquad x \to \infty,$$

where $\alpha = \alpha(\lambda)$ is the unique solution to the equation $E(\lambda Z^2)^{(u/2)} = 1$. In [18] the exact value of c_α is determined.

We observe in particular that $EX^4 = \infty$ if $\lambda^2 > 1/3$ and $EX^2 = \infty$ if $\lambda > 1$. ARCH processes and their various generalizations are examples of stochastic volatility models. As such they are often considered as models for financial log-returns and for exchange rates. There is empirical evidence that log-returns may have infinite fourth moments; see, for example, [17], [19] and [26]. See Figure 2 for an illustration with exchange rate data. This is also confirmed by empirical work performed by the Olsen and Associates Research Group (Zürich) for the foreign exchange rate and interbank market of cash interest rates; see, for example, [7], [20], [28]. Moreover, in the case $EX^2 < \infty$, the ARCH(1) process has vanishing autocorrelations at all nonzero lags. This is in accordance with many log-differenced return series whose sample auto-

TABLE 1
Values of $\alpha = \alpha(\lambda)$ for $\lambda \in (0, 2e^E)$

λ	0.1	0.3	**0.3125**	0.4	0.5	**0.5773**	0.6	0.7	0.8	0.9	**1.0**
α	26.48	8.36	**8.00**	6.09	4.74	**4.00**	3.82	3.17	2.68	2.30	**2.00**
λ	**1.5700**	2.0	2.5	3.0	3.5						
α	**1.00**	0.62	0.34	0.15	0.014						

correlations vanish at almost all lags, with the possible exception of the first lag, but this one is also rather small. See [33] for an empirical study of this phenomenon, which is based on a large variety of financial time series. A further claim is that log-returns have the property of long memory or long-range dependence; see, for example, [5] and [6]. This is partly based on the empirical fact that the absolute values and the squares of those time series may have sample autocorrelations different from zero, even at large lags, and that the ACF decays to zero quite slowly. See Figures 1 and 3–5 for an illustration with real-life and simulated data.

For these reasons [infinite fourth moment of X_t, sample ACF of the X_t not significantly different from 0, and the sample ACF's of $(|X_t|)$ and (X_t^2) decay slowly], it is of interest to study the sample autocovariances and autocorrelations of those processes in order to identify what, if anything, these quantities estimate. For example, if $EX^4 = \infty$, the ergodic theorem does not apply for the partial sums of the X_t^4 and $X_t^2 X_{t+h}^2$, and therefore it is not a priori clear what the sample ACF of the X_t^2 estimates. Similar remarks apply if $EX^2 = \infty$ and one considers the sample ACF of (X_t) or $(|X_t|)$. There is empirical evidence from simulated squared ARCH data showing that the sample autocorrelations do not converge (in probability or a.s.) to a constant, that is, different simulated sample paths from the same model yield completely different values of the sample autocorrelations, even if the sample size is huge. See Figure 4 for an illustration. In what follows we give a rigorous asymptotic theory for the sample autocovariances and autocorrelations of an ARCH(1) process, which explains this phenomenon.

Before embarking on a derivation of the asymptotic theory of the sample ACF of an ARCH process, we first establish the fundamental convergence result for the point processes based on the lagged process.

THEOREM 4.1. *Let (X_t) be the ARCH(1) process defined by* (4.1), *and for fixed $m \geq 0$ set* $\mathbf{X}_t = (X_t, \ldots, X_{t+m})$. *Let (a_n) be a sequence of constants such that* (2.2) *holds. Then the conditions of Theorem* 2.8 *are met, and hence*

$$N_n = \sum_{t=1}^{n} \varepsilon_{\mathbf{X}_t / a_n} \to_d N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}},$$

where (P_i) and $(\mathbf{Q}_{ij})$ are defined in the statement of the theorem.

PROOF. The joint regular variation of all finite-dimensional distributions follows from Lemma A.1. In addition, the process is strongly mixing (see Lemma A.3) so that the mixing condition $\mathscr{A}(a_n)$ is also met. Next, we verify that (2.10) holds and that γ as defined in (2.11) is positive. The latter follows an argument along the lines given for Lemma A.1. For $t > 0$, we have

$$X_t^2 = X_0^2 \prod_{j=1}^{t} \left(\lambda Z_j^2 \right) + (\beta / \lambda) \sum_{j=1}^{t} \prod_{m=j}^{t} \left(\lambda Z_m^2 \right) = X_0^2 I_{t,1} + I_{t,2},$$
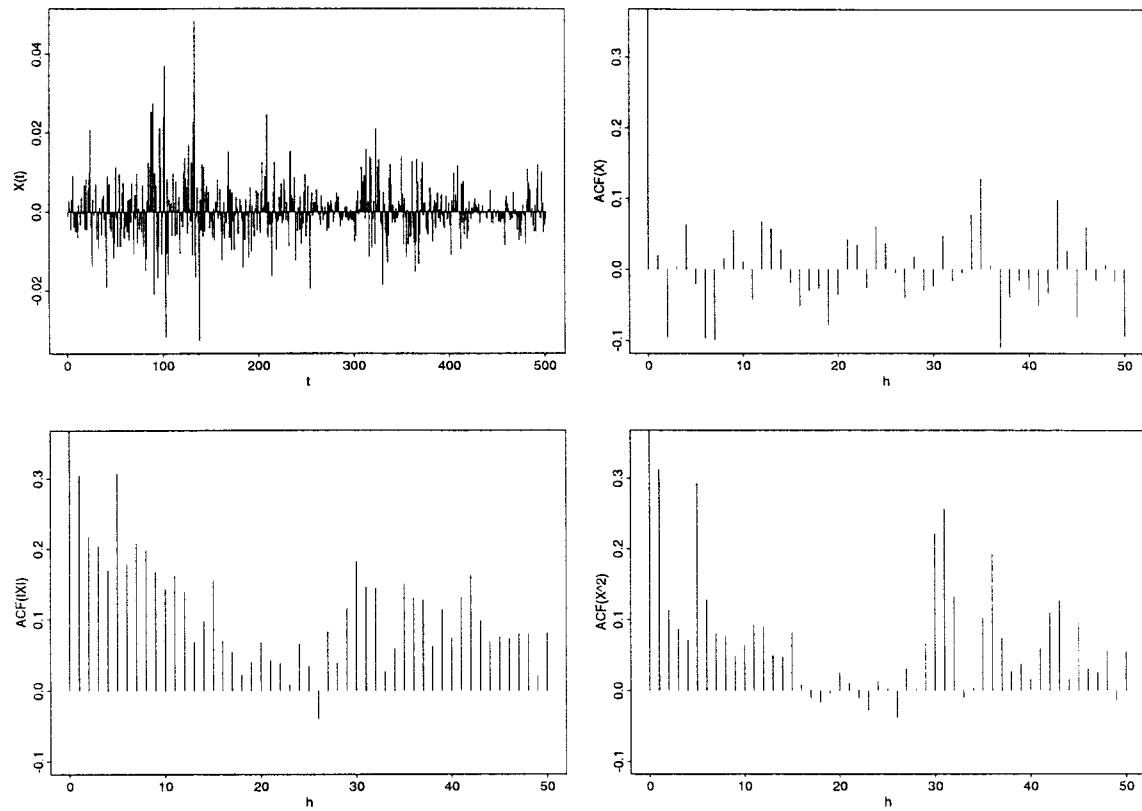
FIG. 1. *Top*: 500 *daily log-returns of the NZ/US $ exchange rate* (*left*) *and the sample ACF of the data* (*right*). *Bottom*: *the sample ACF of the absolute values of the data* (*left*) *and the sample ACF of the squared data* (*right*).
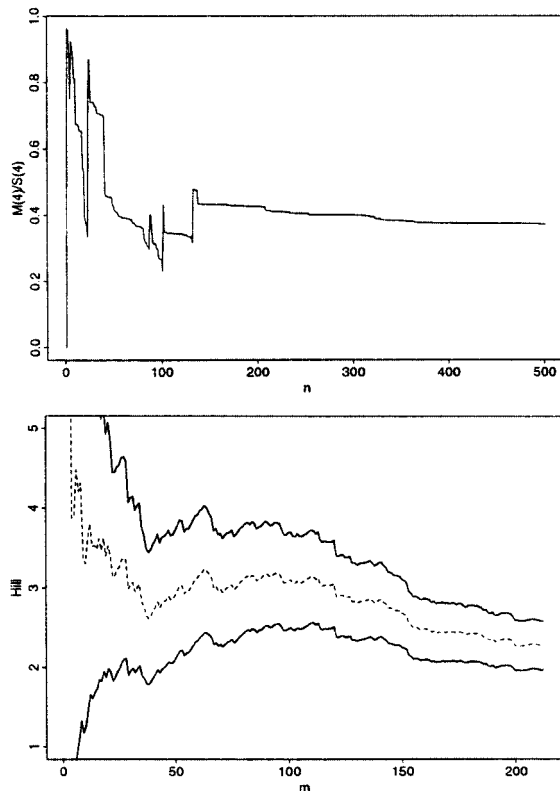
FIG. 2. *Empirical evidence that the exchange rate log-returns do not come from a distribution with finite fourth moment. Left*: *the ratio of maximum* $M_n(4)$ *and sum* $S_n(4)$ *for the fourth power of the data. If* $EX^4 < \infty$, $M_n(4)/S_n(4) \to 0$ *a.s. Right*: *the Hill plot of the tail index* α *based on the assumption that* $P(X > x) = x^{-\alpha}L(x)$ *for some slowly varying* L *and asymptotic confidence bands. The estimate is based on* m (*horizontal axis*) *upper order statistics of a sample of* 1,082 *data. See* [17], *Chapter* 6, *for more background information on tail estimation.*

and hence

$$
(4.4) \quad
\begin{aligned}
P\big(X_t^2 > a_n^2 y | X_0^2 > a_n^2 y\big) \\
\leq P\big(X_0^2 I_{t,1} > a_n^2 y/2 | X_0^2 > a_n^2 y\big) + P\big(I_{t,2} > a_n^2 y/2 | X_0^2 > a_n^2 y\big).
\end{aligned}
$$

Choose $\delta \in (0, \alpha/2)$. Then, using Markov's inequality and Karamata's theorem, the $\limsup_{n \to \infty}$ of the first term on the right of (4.4) is bounded above by

$$
\limsup_{n \to \infty} EI_{t,1}^{\delta}(2/y)^{\delta} a_n^{-2\delta} E\big(|X_0|^{2\delta} I_{\{X_0^2 > a_n^2 y\}}\big) / P\big(X_0^2 > a_n^2 y\big) \leq Cb^t,
$$

where, by the choice of δ, $b = E(\lambda Z^2)^{\delta} < 1$ and C is a constant independent of t. As for the second term in (4.4), we have

$$
I_{t,2} =_d (\beta/\lambda) \sum_{j=1}^{t} \prod_{m=1}^{j} \big(\lambda Z_m^2\big) \uparrow Y \quad \text{a.s.,}
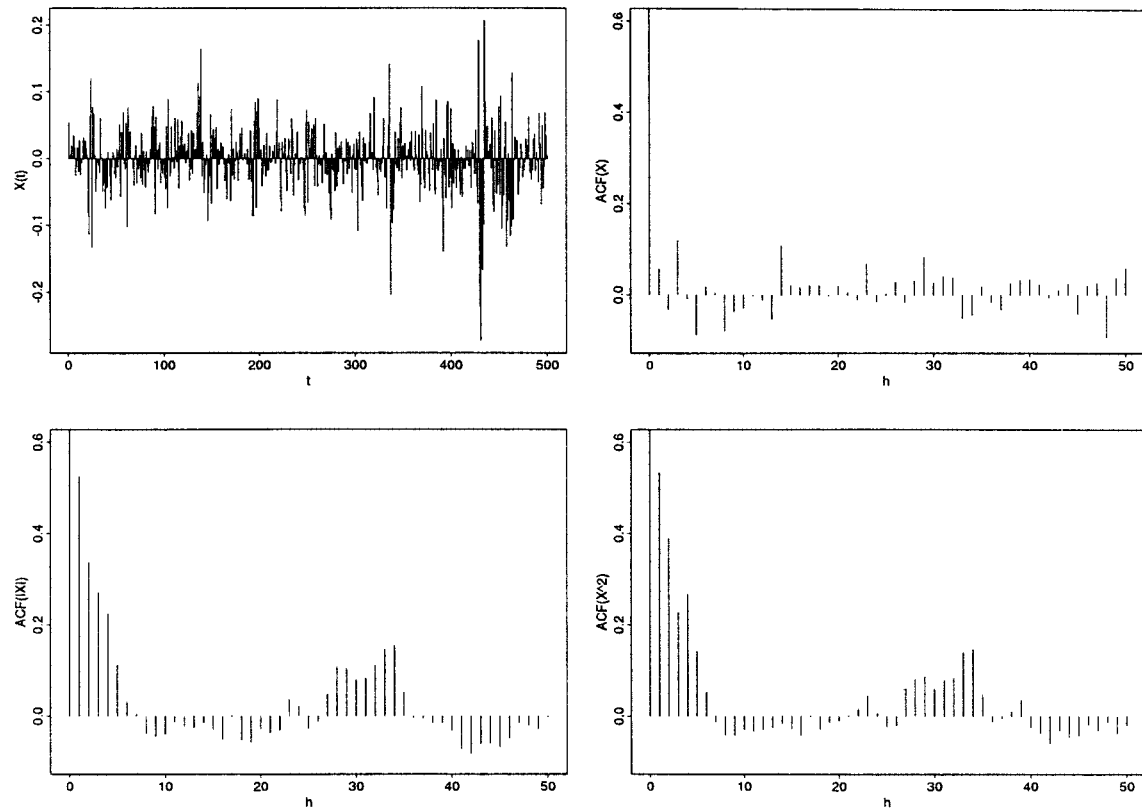$$

FIG. 3. *Top*: 500 *realizations of the ARCH*(1) *process* $X_t = (0.001 + 0.6X_{t-1})^{1/2}Z_t$ (*left*) *and the sample ACF of the data* (*right*). *Bottom*: *the sample ACF of the absolute values of the data* (*left*) *and the sample ACF of the squared data* (*right*).

$Y =_d X_0^2$ [see (4.2)] and since $I_{t,2}$ is independent of X_0,

$$P\big(I_{t,2} > a_n^2 y/2 | X_0^2 > a_n^2 y\big) \leq P\big(Y > a_n^2 y/2\big) = O(n^{-1}).$$

Combining these two bounds for the terms in (4.4), it follows that

$$\lim_{k \to \infty} \limsup_{n \to \infty} P\left(\bigvee_{k \leq |t| \leq r_n} |\mathbf{X}_t| > a_n y \,\middle|\, |\mathbf{X}_0| > a_n y\right)$$

$$\leq \lim_{k \to \infty} \limsup_{n \to \infty} 2(m+1) \sum_{t=k}^{r_n+m} P\big(X_t^2 > a_n^2 y^2 | X_0^2 > a_n^2 y^2\big)$$

$$\times P\big(X_0^2 > a_n^2 y^2\big)/P(|\mathbf{X}_0| > a_n y)$$

$$\leq \lim_{k \to \infty} (\text{const}) \sum_{t=k}^{\infty} b^t$$

$$= 0.$$

This completes the verification of (2.10). $\square$

REMARK 4.2. In the special case when $m = 0$ in Theorem 4.2, we obtain

$$\sum_{t=1}^{n} \varepsilon_{X_t/a_n} \to_d N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i Q_{ij}},$$

which is an extension of Theorem 2.1 in [15] on the convergence of the time-normalized point process of exceedances. In this case the extremal index of (X_t^2) can be calculated as follows. For $k > 0$ fixed, define $\mathbf{Y} = (1, \lambda^{1/2} Z_1, \ldots, \lambda^{k+1/2}(Z_1 \cdots Z_{2k+1}))$. Applying Lemma A.1(ii) with $h = 0$ and $m = 2k$ and using the independence of Y_{k+1} with $(Y_j/Y_{k+1}, j = k+2, \ldots, 2k+1)$, we find that

$$E\left(|\theta_0^{(k)}|^\alpha - \bigvee_{j=1}^{k} |\theta_j^{(k)}|^\alpha\right)_+ \Bigg/ E|\theta_0^{(k)}|^\alpha$$

$$= E\left(|Y_{k+1}|^\alpha \left(1 - \bigvee_{j=k+2}^{2k+1} |Y_j/Y_{k+1}|^\alpha\right)_+\right) \Bigg/ E|Y_{k+1}|^\alpha$$

$$= E\left(1 - \bigvee_{j=k+2}^{2k+1} |Y_j/Y_{k+1}|^\alpha\right)_+$$

$$= E\left(1 - \bigvee_{j=1}^{k} A_j^\alpha\right)_+ ,$$

where $A_j = \prod_{i=1}^{j} |\lambda^{1/2} Z_i|$. The last expectation can be rewritten as

$$\int_1^\infty P\left(\bigvee_{j=1}^{k} A_j \le y^{-1}\right) \alpha y^{-\alpha-1}\, dy,$$

which converges to

$$\gamma = \int_1^\infty P\left(\bigvee_{j=1}^{\infty} A_j \le y^{-1}\right) \alpha y^{-\alpha-1}\, dy,$$

as $k \to \infty$. This expression agrees with the value of the extremal index of the squared process given in equation (2.2) of [15].

A. *Sample* ACVF *and* ACF *of* X_t. We derive the limiting behavior of $\gamma_{n,X}(h)$ and $\rho_{n,X}(h)$ for three cases: (1) $\alpha \in (0,2)$, (2) $\alpha \in (2,4)$ and (3) $\alpha \in (4, \infty)$. In the first case, the variance of X is infinite and $\rho_{n,X}(h)$ has a random limit without any normalization. For $\alpha \in (2,4)$, the process has a finite variance but infinite fourth moment and

$$n a_n^{-2} \rho_{n,X}(h) \to_d S_h,$$

where S_h has an $\alpha/2$-stable distribution. Finally, if $\alpha > 4$, then the process has a finite fourth moment and $n^{1/2} \rho_{n,X}(h)$ is asymptotically normal.

A(1). CASE $\alpha \in (0,2)$. Since the point process result of Theorem 4.2 holds, a direct application of Theorem 3.5(i) yields

$$\left(n a_n^{-2}\left(\gamma_{n,X}(h)\right)\right)_{h=0,\ldots,m} \to_d (V_h)_{h=0,\ldots,m}$$

and

$$\left(\rho_{n,X}(h)\right)_{h=1,\ldots,m} \to_d (V_h/V_0)_{h=1,\ldots,m},$$

where V_h is defined in the theorem.

A(2). CASE $\alpha \in (2,4)$. First, note that by the symmetry of the distribution of Z_t, the random variables for $h > 0$

$$X_t X_{t+h} I_{\{|X_t X_{t+h}| \le a_n^2 \epsilon\}} \quad \text{and} \quad X_s X_{s+h} I_{\{|X_s X_{s+h}| \le a_n^2 \epsilon\}}$$

have mean 0 and are uncorrelated for $s \ne t$. Using this property in conjunction with the regular variation of $X_0 X_h$ [with tail index $\alpha/2$; cf. Lemma A.1(ii)], we have

$$\operatorname{var}\left(a_n^{-2} \sum_{t=1}^{n-h} X_t X_{t+h} I_{\{|X_t X_{t+h}| \le a_n^2 \epsilon\}}\right)$$

$$= (n-h) a_n^{-4} E X_0^2 X_h^2 I_{\{|X_0 X_h| \le a_n^2 \epsilon\}}$$

$$\sim n a_n^{-4} \alpha (4-\alpha)^{-1} \left(a_n^2 \epsilon\right)^2 P\left(|X_0 X_h| > a_n^2 \epsilon\right)$$

$$\to \text{const } \epsilon^{2-\alpha/2} \quad \text{as } n \to \infty,$$

$$\to 0 \quad \text{as } \epsilon \to 0,$$

where the asymptotic relation follows from Karamata's theorem. Thus, condition (3.4) of Theorem 3.5 is met for $h = 1, \ldots, m$ and since $\gamma_X(h) = 0$ for $h \geq 1$, we conclude

$$\left(na_n^{-2}(\gamma_{n,X}(h))\right)_{h=1,\ldots,m} \to_d (V_h)_{h=1,\ldots,m},$$

where V_h is defined in the theorem. In addition, an application of the ergodic theorem yields $\gamma_{n,X}(0) \to_P \gamma_X(0)$ so that

$$\left(na_n^{-2}\rho_{n,X}(h)\right)_{h=1,\ldots,m} \to_d \gamma_X^{-1}(0)(V_h)_{h=1,\ldots,m}.$$

A(3). CASE $\alpha \in (4, \infty)$. Since the stationary multivariate time series $(X_t^2, X_t X_{t+1}, \ldots, X_t X_{t+m})$ has a finite $(2 + \delta)$th moment $(\delta > 0)$ and is strongly mixing with a geometric rate (Lemma A.3), it follows from standard limit theorems for strongly mixing sequences (see, for example, Theorem 18.5.3 of Ibragimov and Linnik [21]) that

$$\left(n^{1/2}(\gamma_{n,X}(h) - \gamma_X(h))\right)_{h=0,\ldots,m} \to_d (G_h)_{h=0,\ldots,m},$$

where the limit vector has a multivariate normal distribution with mean zero. Applying the argument used to establish (3.6), we conclude that

$$\left(n^{1/2}(\rho_{n,X}(h))\right)_{h=1,\ldots,m} \to_d \gamma_X^{-1}(0)(G_h)_{h=1,\ldots,m}.$$

Notice that the sequences of scaling constants which ensure nondegenerate convergence of $\rho_{n,X}(h)$ for the various cases are of the form

$$\begin{cases} 1, & \text{if } \alpha \in (0,2), \\ \text{const } n^{-1/2\alpha}, & \text{if } \alpha \in (2,4), \\ \text{const } n^{1/2}, & \text{if } \alpha \in (4,\infty). \end{cases}$$

This is in contrast to the linear process case with regularly varying noise variables, where the scaling increases with decreasing α; compare [11, 12].

B. *Sample* ACVF *and* ACF *of* X_t^2. The sample ACVF and ACF based on the squares of the process are defined by

$$\gamma_{n,X^2}(h) = n^{-1} \sum_{t=1}^{n-h} X_t^2 X_{t+h}^2, \qquad h \geq 0,$$

and

$$\rho_{n,X^2}(h) = \gamma_{n,X^2}(h)/\gamma_{n,X^2}(0), \qquad h \geq 0,$$

respectively. Like the situation considered above for $\gamma_{n,X}$ and $\rho_{n,X}$, there are three intervals of α that give rise to vastly different limit behavior of γ_{n,X^2} and ρ_{n,X^2}. These regions are determined by the existence or nonexistence of the second and fourth moments of X_t^2.
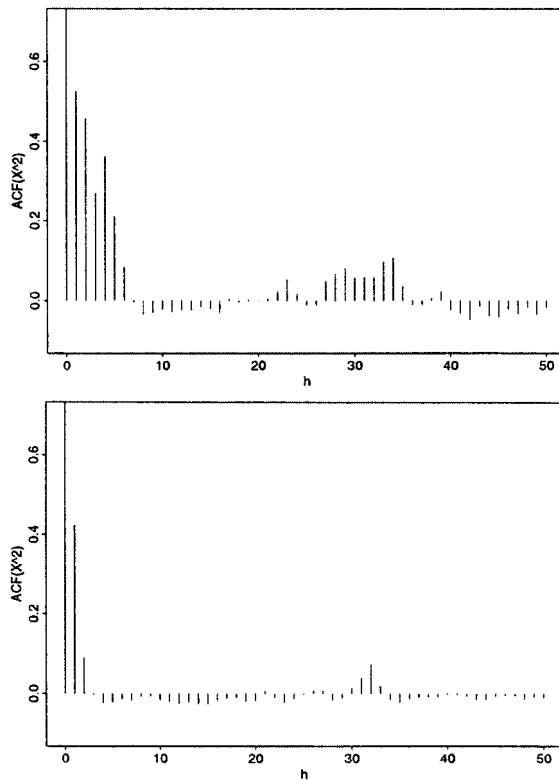
FIG. 4. *Two paths of the sample ACF of the squares of* 500 *different realizations of the ARCH*(1) *process* $X_t = (0.001 + 0.7X_{t-1})^{1/2}Z_t$. *These simulations illustrate that the sample ACF do not converge to constants.*

B(1). CASE $\alpha \in (0, 4)$. It is easy to see that the point process convergence in Theorem 4.2 remains valid for the squared process, that is,

$$(4.5) \qquad N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-2}(X_t^2, \ldots, X_{t+m}^2)} \to_d N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i^2 \tilde{\mathbf{Q}}_{ij}},$$

where $\tilde{Q}_{ij}^{(h)} = (Q_{ij}^{(h)})^2$, $h = 0, \ldots, m$. Since $(X_t^2, \ldots, X_{t+m}^2)$ is jointly regularly varying with tail index $\alpha/2$, it follows from Theorem 3.5(i) that

$$\left(na_n^{-4}(\gamma_{n, X^2}(h))\right)_{h=0,\ldots,m} \to_d (V_h)_{h=0,\ldots,m},$$

and

$$\left(\rho_{n, X^2}(h)\right)_{h=1,\ldots,m} \to_d (V_h/V_0)_{h=1,\ldots,m},$$

where

$$V_h = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^4 \left(Q_{ij}^{(0)}\right)^2 \left(Q_{ij}^{(h)}\right)^2, \qquad h = 0, \ldots, m.$$

B(2). CASE $\alpha \in (4, 8)$. Due to the difficulty of verifying condition (3.4) directly, we establish convergence in distribution of the sample ACVF directly from the point process convergence in (4.5). We only establish joint convergence of $(\gamma_{n, X^2}(0), \gamma_{n, X^2}(1))$, since the extension to arbitrary lags is a straightforward generalization. Now using the representation $X_{t+1}^4 = (\beta + \lambda X_t^2)^2 Z_{t+1}^4$, write

$$a_n^{-4} \sum_{t=1}^n \left(X_{t+1}^4 - EX^4 \right)$$

$$= a_n^{-4} \sum_{t=1}^n \left(\beta + \lambda X_t^2 \right)^2 (Z_{t+1}^4 - 3) I_{\{|X_t| \leq a_n \epsilon\}}$$

(4.6)

$$+ a_n^{-4} \sum_{t=1}^n \left(\beta + \lambda X_t^2 \right)^2 (Z_{t+1}^4 - 3) I_{\{|X_t| > a_n \epsilon\}}$$

$$+ 3a_n^{-4} \sum_{t=1}^n \left(\left(\beta + \lambda X_t^2 \right)^2 - E\left(\beta + \lambda X^2 \right)^2 \right)$$

$$= \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

Since the summands comprising I are uncorrelated, we have, by Karamata's theorem,

$$\mathrm{var}(\mathrm{I}) = na_n^{-8} \, \mathrm{var}\!\left(\left(\beta + \lambda X_0^2 \right)^2 (Z_1^4 - 3) I_{\{|X_0| \leq a_n \epsilon\}} \right)$$

$$\leq \mathrm{const}\, na_n^{-8} EX^8 I_{\{|X| \leq a_n \epsilon\}}$$

(4.7)

$$\sim \mathrm{const}\, na_n^{-8} (a_n \epsilon)^8 P(|X| > a_n \epsilon)$$

$$\to \mathrm{const}\, \epsilon^{8 - \alpha} \quad \text{as } n \to \infty$$

$$\to 0 \quad \text{as } \epsilon \to 0.$$

As for II, let $\mathbf{x}_t = (x_t^{(0)}, \ldots, x_t^{(m)}) \in \overline{\mathbb{R}}^{m+1} \setminus \{\mathbf{0}\}$ and define the mappings

$$T_{h, \epsilon} : \mathcal{M} \to \overline{\mathbb{R}}$$

by

$$T_{0, \epsilon}\!\left(\sum_{i=1}^\infty n_i \varepsilon_{\mathbf{x}_i} \right) = \sum_{i=1}^\infty n_i \left(x_i^{(0)} \right)^2 I_{\{|x_i^{(0)}| > \epsilon\}},$$

$$T_{1, \epsilon}\!\left(\sum_{i=1}^\infty n_i \varepsilon_{\mathbf{x}_i} \right) = \sum_{i=1}^\infty n_i \left(x_i^{(1)} \right)^2 I_{\{|x_i^{(0)}| > \epsilon\}},$$

$$T_{h, \epsilon}\!\left(\sum_{i=1}^\infty n_i \varepsilon_{\mathbf{x}_i} \right) = \sum_{i=1}^\infty n_i x_i^{(0)} x_i^{(h-1)} I_{\{|x_i^{(0)}| > \epsilon\}}, \qquad h \geq 2.$$

Since the set $\{\mathbf{x} \in \overline{\mathbb{R}}^2 \setminus \{\mathbf{0}\} : |x_0| > \epsilon\}$ is bounded, the convergence in (4.5) implies

(4.8)

$$\mathrm{II} = T_{1, \epsilon} N_n - 3\lambda^2 T_{0, \epsilon} N_n + o_P(1)$$

$$\to_d T_{1, \epsilon} N - 3\lambda^2 T_{0, \epsilon} N =: S(\epsilon, \infty).$$

Note that $ES(\epsilon, \infty) = 0$. Using (4.7) and the argument given in [8], pages 897 and 898, $S(\epsilon, \infty) \to_d V_0^*$, say, as $\epsilon \to 0$.

Turning to III, it is easy to see that

$$\text{III} = 3a_n^{-4} \sum_{t=1}^{n} \lambda^2 \big(X_{t+1}^4 - EX^4 \big) + o_P(1),$$

so that from (4.6)–(4.8), we have

$$(4.9) \qquad a_n^{-4} \sum_{t=1}^{n} \big(X_{t+1}^4 - EX^4 \big) \to_d (1 - 3\lambda^2)^{-1} V_0^* =: V_0.$$

As for $\gamma_{n, X^2}(1)$, we proceed as above and write

$$a_n^{-4} \sum_{t=1}^{n} \big(X_t^2 X_{t+1}^2 - \gamma_{X^2}(1) \big)$$

$$= a_n^{-4} \sum_{t=1}^{n} X_t^2 \big(\beta + \lambda X_t^2 \big) \big(Z_{t+1}^2 - 1 \big) I_{\{|X_t| \le a_n \epsilon\}}$$

$$+ a_n^{-4} \sum_{t=1}^{n} X_t^2 \big(\beta + \lambda X_t^2 \big) \big(Z_{t+1}^2 - 1 \big) I_{\{|X_t| > a_n \epsilon\}}$$

$$+ a_n^{-4} \sum_{t=1}^{n} \big(X_t^2 \big(\beta + \lambda X_t^2 \big) - \gamma_{X^2}(1) \big).$$

Since $\gamma_{X^2}(1) = \beta EX^2 + \lambda EX^4$, it follows from (4.9) and the argument given above that

$$na_n^{-4} \big(\gamma_{n, X^2}(1) - \gamma_{X^2}(1) \big) \to_d V_1^* + \lambda V_0 =: V_1,$$

where V_1^* is the distributional limit of $T_{2, \epsilon} N - \lambda T_{1, \epsilon} N$ as $\epsilon \to 0$. In general,

$$\big(na_n^{-4} \big(\gamma_{n, X^2}(h) - \gamma_{X^2}(h) \big) \big)_{h=0, \ldots, m} \to_d (V_h)_{h=0, \ldots, m},$$

where

$$V_0 = V_0^* (1 - 3\lambda^2)^{-1}, \qquad V_h = V_h^* + \lambda V_{h-1}, \, h \ge 1,$$

and $(V_0^*, \ldots, V_m^*)$ is the distributional limit of $(T_{1, \epsilon} N - 3\lambda^2 T_{0, \epsilon} N, (T_{h+1, \epsilon} N - \lambda T_{h, \epsilon} N)_{h=1, \ldots, m})$ as $\epsilon \to 0$. Moreover, the Taylor series argument used earlier gives

$$\big(na_n^{-4} \big(\rho_{n, X^2}(h) - \rho_{X^2}(h) \big) \big)_{h=1, \ldots, m} \to_d \gamma_{X^2}^{-1}(0) \big(V_h - \rho_{X^2}(h) V_0 \big)_{h=1, \ldots, m}.$$

B(3). CASE $\alpha \in (8, \infty)$. The sample ACVF and ACF are asymptotically normal in this case. The proof is identical to that described in part A(3) above.

C. *Sample* ACVF *and* ACF *of* $|X_t|$. The asymptotic behavior of the sample ACVF and ACF based on the absolute values of the process is similar to the situation considered in part A. The proofs of the $\alpha \in (0, 2)$ and $\alpha \in (4, \infty)$ cases are identical to those given for parts A(1) and A(3), respec-
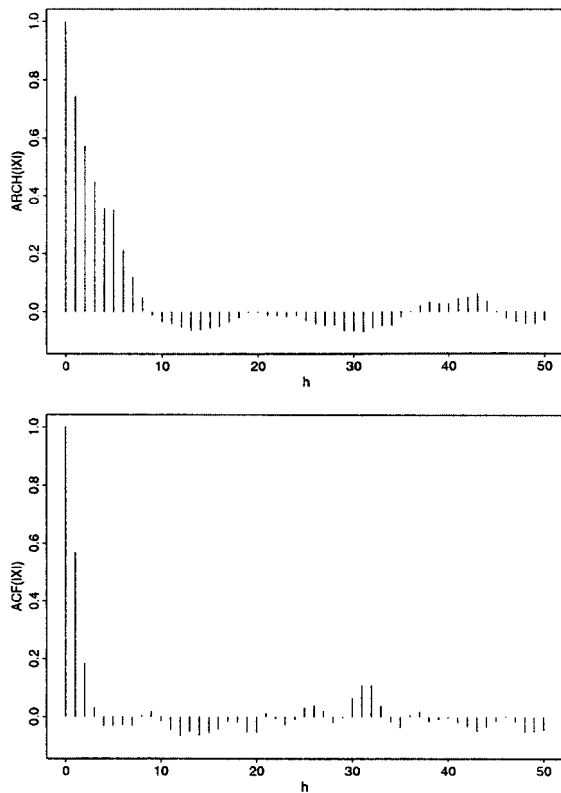
FIG. 5. *Two paths of the sample ACF of the absolute values of* 500 *different realizations of the ARCH*(1) *process* $X_t = (0.001 + X_{t-1})^{1/2} Z_t$. *These simulations illustrate that the sample ACF do not converge to constants.*

tively, while the proof in the $\alpha \in (2, 4)$ case follows the argument given for part B(2). The relevant point process convergence for dealing with the absolute values is now

$$(4.10) \qquad N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1}(|X_t|, \ldots, |X_{t+m}|)} \to_d N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i(|Q_{ij}^{(0)}|, \ldots, |Q_{ij}^{(m)}|)},$$

where (P_i) and $\mathbf{Q}_{ij} = (Q_{ij}^{(0)}, \ldots, Q_{ij}^{(m)})$ are specified in Theorem 4.2.

C(1). CASE $\alpha \in (0, 2)$. We have, from a direct application of Theorem 3.5(i), that

$$\left(na_n^{-4}\gamma_{n,|X|}(h)\right)_{h=0,\ldots,m} \to_d (V_h)_{h=0,\ldots,m}$$

and

$$\left(\rho_{n,|X|}(h)\right)_{h=1,\ldots,m} \to_d (V_h/V_0)_{h=1,\ldots,m},$$

where

$$V_h = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 \big(|Q_{ij}^{(0)} Q_{ij}^{(h)}| \big), \qquad h = 0, \ldots, m.$$

C(2). CASE $\alpha \in (2, 4)$. As in the case considered for part B(2), we only provide a proof of the joint convergence of $(\gamma_{n,|X|}(0), \gamma_{n,|X|}(1))$. Using the argument given in B(2) for the sample fourth moment, we have

$$a_n^{-2} \sum_{t=1}^n \big(X_t^2 - EX^2 \big)$$

$$= a_n^{-2} \sum_{t=1}^n \big(\beta + \lambda X_t^2 \big) \big(Z_{t+1}^2 - 1 \big) + a_n^{-2} \sum_{t=1}^n \lambda \big(X_t^2 - EX^2 \big) + o_P(1).$$

It then follows that

(4.11) $$a_n^{-2} \sum_{t=1}^n \big(X_t^2 - EX^2 \big) \to_d (1 - \lambda)^{-1} V_0^* =: V_0,$$

where V_0^* is the distributional limit of $T_{1,\epsilon} N - T_{0,\epsilon} N$ as $\epsilon \to 0$ and the mappings $T_{1,\epsilon}$ and $T_{0,\epsilon}$ are defined in part B(2).

For the lag one covariances, we begin with the decomposition,

$$a_n^{-2} \sum_{t=1}^n \big(|X_t X_{t+1}| - E|X_1 X_2| \big)$$

$$= a_n^{-2} \sum_{t=1}^n |X_t| \big(\beta + \lambda X_t^2 \big)^{1/2} \big(|Z_{t+1}| - \pi^{-1} \big) + a_n^{-2} \pi^{-1} \sum_{t=1}^n \lambda^{1/2} \big(X_t^2 - EX^2 \big)$$

$$+ a_n^{-2} \pi^{-1} \sum_{t=1}^n \lambda^{1/2} \big(f(X_t) - Ef(X) \big)$$

$$= \mathrm{I} + \mathrm{II} + \mathrm{III},$$

where $f(x) = |x| \big((\beta + \lambda x^2)^{1/2} - \lambda^{1/2} |x| \big)$. Using the bound $0 \le f(x) \le \beta^{1/2} |x|$, we note that $f^2(x) \le \beta x^2 \le \beta V(x)$, where $V(x) = 1 + |x|^\delta$ with $\delta \in (2, \alpha)$. Applying the result of Remark A.4, we obtain

$$\mathrm{var}(\mathrm{III}) = a_n^{-4} \sum_{s=1}^n \sum_{t=1}^n \mathrm{cov} \big(f(X_s), f(X_t) \big)$$

$$\le n a_n^{-4} 2\beta (\mathrm{const}) \sum_{h=0}^n a^h$$

$$\to 0,$$

as $n \to \infty$, since $n a_n^{-4} = O(n^{1 - 4/\alpha}) = o(1)$.

Next we consider the term I. We first show

$$a_n^{-2} \sum_{t=1}^{n} |X_t|(\beta + \lambda X_t^2)^{1/2}(|Z_{t+1}| - \pi^{-1})I_{\{|X_t| > a_n\epsilon\}}$$

(4.12)
$$= a_n^{-2} \sum_{t=1}^{n} |X_t X_{t+1}| I_{\{|X_t| > a_n\epsilon\}}$$

$$- \pi^{-1}\lambda^{1/2} a_n^{-2} \sum_{t=1}^{n} X_t^2 I_{\{|X_t| > a_n\epsilon\}} + o_P(1);$$

(4.13)
$$a_n^{-2} \sum_{t=1}^{n} |X_t X_{t+1}| I_{\{|X_t| > a_n\epsilon\}} - \pi^{-1}\lambda^{1/2} a_n^{-2} \sum_{t=1}^{n} X_t^2 I_{\{|X_t| > a_n\epsilon\}}$$

$$\to_d S_\epsilon N - \pi^{-1}\lambda^{1/2} T_{0,\epsilon} N,$$

where $S_\epsilon: \mathcal{M} \to \overline{\mathbb{R}}$ is the mapping $S_\epsilon(\sum_{i=1}^{\infty} n_i \varepsilon_{\mathbf{x}_i}) = \sum_{i=1}^{\infty} n_i |x_i^{(0)} x_i^{(1)}| I_{\{|x_i^{(0)}| > \epsilon\}}$;

(4.14)
$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \mathrm{var}\left(a_n^{-2} \sum_{t=1}^{n} |X_t|(\beta + \lambda X_t^2)^{1/2} \right.$$

$$\left. \times (|Z_{t+1}| - \pi^{-1}) I_{\{|X_t| \le a_n\epsilon\}} \right) = 0;$$

and

(4.15)
$$S_\epsilon N - \pi^{-1}\lambda^{1/2} T_{0,\epsilon} N \to_d V_1^*,$$

as $\epsilon \to 0$. The expectation of the absolute value of the difference between the term on the left and the sum of the first two terms on the right of (4.12) is bounded above by $\pi^{-1}\beta^{1/2} n a_n^{-2} E|X| I_{\{|X| > a_n\epsilon\}}$) which converges to zero by Karamata's theorem. This proves (4.12). The limit in (4.13) is immediate from the point process result in (4.10). The summands in the expression for (4.14) are uncorrelated so that the result holds by an application of Karamata's theorem; compare (4.7). The convergence in (4.15) can be established using (4.14) and the argument given in [8], pages 897 and 898. Combining (4.11)–(4.15), we conclude that $I \to_d V_1^*$. It now follows that

$$\left(n a_n^{-2}(\gamma_{n,|X|}(0) - \gamma_{|X|}(0)), n a_n^{-2}(\gamma_{n,|X|}(1) - \gamma_{|X|}(1)) \right)$$
$$\to_d \left(V_0, V_1^* + \pi^{-1}\lambda^{1/2} V_0 \right)$$

and

$$n a_n^{-2}(\rho_{n,|X|}(1) - \rho_{|X|}(1)) \to_d \gamma_{|X|}^{-1}(0)\left(V_1^* + \pi^{-1}\lambda^{1/2} V_0 - \rho_{|X|}(1) V_0 \right).$$

C(3). CASE $\alpha \in (4, \infty)$. As in Case A(3), $n^{1/2}(\gamma_{n,|X|}(h) - \gamma_{|X|}(h))_{h=0,\ldots,m}$ and $n^{1/2}(\rho_{n,|X|}(h) - \rho_{|X|}(h))_{h=1,\ldots,m}$ are asymptotically normal.

APPENDIX

In what follows, we collect some of the properties of an ARCH(1) process; see (4.1). We use the notation introduced in the previous sections.

Regular variation of the ARCH process.

LEMMA A.1.   *Let (X_t) be a stationary ARCH(1) process with tail index α; see (4.3).*

(i) *Define for $i = 1, 2$, $\mathbf{X}_t^{(i, h)} = (|X_t|^i, |X_{t+h}|^i)$ for some $h \geq 0$ and*

$$\mathbf{X}_t^{(i, h)}(m) = \big(\mathbf{X}_t^{(i, h)}, \dots, \mathbf{X}_{t+m}^{(i, h)}\big), \qquad m \geq 0.$$

Then the random vectors $\mathbf{X}_t^{(i, h)}(m)$ are regularly varying with index α/i; that is, there exists a sequence (x_n) of positive numbers such that

(A.1)
$$nP\left(|\mathbf{X}_t^{(i, h)}(m)| \geq sx_n, \frac{\mathbf{X}_t^{(i, h)}(m)}{|\mathbf{X}_t^{(i, h)}(m)|} \in \cdot\right)$$

$$\rightarrow_v s^{-\alpha/i} \frac{E|\mathbf{Y}^{(i, h)}(m)|^{\alpha/i} I.(\mathbf{Y}^{(i, h)}(m)/|\mathbf{Y}^{(i, h)}(m)|)}{E|\mathbf{Y}^{(i, h)}(m)|^{\alpha/i}},$$

where

$$\mathbf{Y}^{(i, h)}(m)$$
$$= \Big(\big(1, |\lambda^h Z_1^2 \cdots Z_h^2|^{i/2}\big), \dots, \big(|\lambda^m Z_1^2 \cdots Z_m^2|^{i/2}, |\lambda^{m+h} Z_1^2 \cdots Z_{m+h}^2|^{i/2}\big)\Big).$$

(ii) *Define $\mathbf{X}_t^{(h)} = (X_t, X_{t+h})$ for some $h \geq 0$ and*

$$\mathbf{X}_t^{(h)}(m) = \big(\mathbf{X}_t^{(h)}, \dots, \mathbf{X}_{t+m}^{(h)}\big), \qquad m \geq 0.$$

Then the random vectors $\mathbf{X}_t^{(h)}(m)$ are regularly varying with index α; that is, there exists a sequence (x_n) of positive numbers such that

$$nP\left(|\mathbf{X}_t^{(h)}(m)| > sx_n, \frac{\mathbf{X}_t^{(h)}(m)}{|\mathbf{X}_t^{(h)}(m)|} \in \cdot\right)$$

$$\rightarrow_v s^{-\alpha} \frac{E|\mathbf{Y}^{(h)}(m)|^{\alpha} I.(\mathbf{Y}^{(h)}(m)/|\mathbf{Y}^{(h)}(m)|)}{E|\mathbf{Y}^{(h)}(m)|^{\alpha}},$$

where

$$\mathbf{Y}^{(h)}(m) = \Big(\big(1, \lambda^{h/2} Z_1 \cdots Z_h\big), \dots, \big(\lambda^{m/2} Z_1 \cdots Z_m, \lambda^{(m+h)/2} Z_1 \cdots Z_{m+h}\big)\Big).$$

PROOF.   (i) For notational simplicity we restrict ourselves to the case $h = 1$, $i = 2$, that is,

$$\mathbf{X}_t(m) = \mathbf{X}_t^{(2, 1)}(m) = \Big(\big(X_t^2, X_{t+1}^2\big), \dots, \big(X_{t+m}^2, X_{t+m+1}^2\big)\Big).$$

The case $i = 1$ follows in the same way. Iterating the defining equation (4.1), one obtains

$$\mathbf{X}_t(m) = X_t^2\big((1, \lambda Z_{t+1}^2), \ldots, (\lambda^m Z_{t+1}^2 \cdots Z_{t+m}^2, \lambda^{m+1} Z_{t+1}^2 \cdots Z_{t+m+1}^2)\big) + \mathbf{R}_t$$
$$= X_t^2 \mathbf{Y}_t + \mathbf{R}_t,$$

where the remainder $\mathbf{R}_t$ has a tail which does not contribute asymptotically to the tail of $\mathbf{X}_t(m)$. Hence it suffices to show that $X^2 \mathbf{Y} = X_t^2 \mathbf{Y}_t$ is regularly varying. Fix a Borel set $A \subset \mathbb{S}^{2m-1}$. Choose x_n such that $P(X^2 > x_n) \sim n^{-1}/E|\mathbf{Y}|^{\alpha/2}$. A result of [3] (see also [29]) says that if ξ is a nonnegative random variable with a regularly varying tail of index $\tilde{\alpha} > 0$ and if η is another nonnegative random variable independent of ξ and with $E|\eta|^{\kappa} < \infty$ for some $\kappa > \tilde{\alpha}$, then

$$P(\eta \xi > x) \sim E\eta^{\tilde{\alpha}} P(\xi > x), \qquad x \to \infty.$$

An application of the latter result with $\xi = X^2$, $\tilde{\alpha} = \alpha/2$, $\eta = |\mathbf{Y}| I_{\{\mathbf{Y}/|\mathbf{Y}| \in A\}}$ yields the desired relation (A.1).

(ii) This part follows from the regular variation of the absolute values and by the observation that $(X_t) = (|X_t|r_t)$ where the sequence (r_t) is independent of $(|X_t|)$. Alternatively, one can follow the lines of the proof above. $\square$

Strong mixing of the ARCH.

LEMMA A.2.   *Let (X_t) be a stationary ARCH(1) process; see (4.1). Then (X_t) is strongly mixing with geometric rate; that is, there exist constants $K > 0$ and $a \in (0, 1)$ such that*

$$\sup_{A \in \sigma(X_s, s \le 0),\, B \in \sigma(X_s, s > k)} |P(A \cap B) - P(A)P(B)| \le Ka^k.$$

PROOF.   We follow the theory for Markov chains in [27]. The case $\lambda < 1$ can also be derived from [16]. Observe that $X_t = (\beta + \lambda X_{t-1}^2)^{1/2} Z_t$ is a Markov chain with state space $\mathbb{R}$ and transition probabilities $P(x, \cdot)$. The latter correspond to an $N(0, \beta + \lambda x^2)$ distribution. By Theorem 16.1.5 in [27], it suffices to show that the chain is V-uniformly ergodic. For this one has to check the *geometric drift condition* (V4) in [27], page 367: there exist an extended real-valued function $V \colon \mathbb{R} \to [1, \infty]$, a measurable set C, constants $\tilde{\beta} > 0$, $b < \infty$ such that

$$(A.2) \qquad \int V(y) P(x, dy) \le (1 - \tilde{\beta}) V(x) + b I_C(x), \qquad x \in \mathbb{R}.$$

First we show that for any $M > 0$, the set $C = [-M, M]$ is *petite* in the sense of [27], page 121; there exists a non-trivial measure ν on the Borel sets $\mathcal{B}$ of

$\mathbb{R}$ such that $P(x, B) \geq \nu(B)$ for all $x \in C$ and $B \in \mathscr{B}$. For all $x \in C$,

$$P(x, B) = \left(2\pi\left(\beta + \lambda x^2\right)\right)^{-1/2} \int_B e^{-0.5 y^2 / (\beta + \lambda x^2)} \, dy$$

$$\geq \frac{\beta^{1/2}}{\left(\beta + M^2 \lambda\right)^{1/2}} \Phi_\beta(B) =: \nu(B),$$

where $\Phi_\beta(B)$ denotes the $N(0, \beta)$ distribution. Hence C is petite. Now choose $V(x) = 1 + |x|^\delta$ for some $\delta > 0$. We check (A.2) directly. We first note that since the function $h_\lambda(u) = E(\lambda Z^2)^{u/2}$ is convex and satisfies $h_\lambda(0) = h_\lambda(\alpha) = 1$; see [17], Lemma 8.4.6, we have $h_\lambda(u) < 1$ for $u < \alpha$. Now for $\delta < \alpha$

$$\int V(y) P(x, dy) = 1 + \int_{\mathbb{R}} |y|^\delta P(x, dy)$$

$$= 1 + \left(\beta + \lambda x^2\right)^{\delta/2} E|Z|^\delta$$

$$= 1 + \left(\beta \lambda^{-1} + x^2\right)^{\delta/2} E|\lambda Z^2|^{\delta/2}.$$

It follows that there exist $\tilde{\beta} \in (0, 1)$ and M large such that the right-hand side is bounded above by $(1 - \tilde{\beta})V(x)$ for $|x| > M$. Choosing $b = 1 + (\beta + \lambda M^2)^{\delta/2} E|Z|^\delta$ and $C = [-M, M]$, (A.2) now follows. $\square$

The following is a consequence of Lemma A.2, the fact that (X_t) is a Markov chain and that strong mixing is a property of the underlying σ-fields.

LEMMA A.3.   *For a stationary ARCH(1) process (X_t) and every $m \geq 0$, the sequence of random vectors $\mathbf{X}_t = (X_t, \ldots, X_{t+m})$ is strongly mixing with geometric rate. Moreover, for every measurable function f, the stationary sequence $(f(\mathbf{X}_t))$ is strongly mixing with geometric rate.*

REMARK 5.4.   A particular consequence of Lemma 5.3 is that

$$\sup_{f, g} |\operatorname{cov}(f(\mathbf{X}_0), g(\mathbf{X}_n))| \leq \operatorname{const} a^n,$$

for some $a < 1$ and measurable functions f and g with $|f(x)| \leq 1 + |x|^\delta$ and $|g(x)| \leq 1 + |x|^\delta$ for any $\delta \in (0, \alpha/2)$. These covariances are well defined; [27], page 388. The latter condition allows one to apply standard results on the CLT for strongly mixing sequence; see, for example, [21], Chapter 18.

REFERENCES

[1] ANDERSON, P. L. and MEERSCHAERT, M. M. (1997). Periodic moving averages of random variables with regularly varying tails. *Ann. Statist.* **25** 771–785.

[2] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation.* Cambridge Univ. Press.

[3] BREIMAN, L. (1965). On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10** 323–331.

[4] BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.

[5] CHEUNG, Y.-W. (1993). Long memory in foreign-exchange rates. *J. Bus. Econom. Statist.* **11** 93–101.

[6] DACOROGNA, U. A., MÜLLER, U. A., NAGLER, R. J., OLSEN, R. B. and PICTET, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *J. Internat. Money and Finance* **12** 413–438.

[7] DACOROGNA, M. M., MÜLLER, U. A., PICTET, O. V. and DE VRIES, C. G. (1995). The distribution of extremal foreign exchange rate returns in extremely large data sets. Preprint. Olsen and Associates, Zürich.

[8] DAVIS, R. A. and HSING, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** 879–917.

[9] DAVIS, R. A. and MARENGO, J. (1990). Limit theory for the sample covariance and correlation matrix function of a class of multivariate linear processes. *Stochastic Models* **6** 483–498.

[10] DAVIS, R. A., MARENGO, J. and RESNICK, S. I. (1985). Extremal properties of a class of multivariate moving averages. In *Proceedings of the 45th Session of the ISI, Amsterdam* **26.2** 14. Internat. Statist. Inst., Voorburg.

[11] DAVIS, R. A. and RESNICK, S. I. (1985). More limit theory for the sample correlation function of moving averages. *Stochastic Process. Appl.* **20** 257–279.

[12] DAVIS, R. A. and RESNICK, S. I. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14** 533–558.

[13] DAVIS, R. A. and RESNICK, S. I. (1996). Limit theory for bilinear processes with heavy tailed noise. *Ann. Appl. Probab.* **6** 1191–1210.

[14] DE HAAN, L. and RESNICK, S. I. (1981). On the observation closest to the origin. *Stochastic Process. Appl.* **11** 301–308.

[15] DE HAAN, L., RESNICK, S. I., ROOTZÉN, H. and DE VRIES, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32** 213–224.

[16] DIEBOLT, J. and GUÉGAN, D. (1993). Tail behaviour of the stationary density of general non-linear autoregressive processes of order 1. *J. Appl. Probab.* **30** 315–329.

[17] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events in Insurance and Finance.* Springer, Berlin.

[18] GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** 126–166.

[19] GROENENDIJK, P. A., LUCAS, A. L. and DE VRIES, C. G. (1995). A note on the relationship between GARCH and symmetric stable processes. *J. Empirical Finance* **2** 253–264.

[20] GUILLAUME, D. M., DACOROGNA, M. M., DAVÉ, R. R., MÜLLER, U. A., OLSEN, R. B. and PICTET, O. V. (1995). For the bird's eye to the microscope. A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* **1** 95–129.

[21] IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

[22] KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Akademie, Berlin.

[23] KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248.

[24] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[25] LEADBETTER, M. R. and ROOTZÉN, H. (1988). Extremal theory for stochastic processes. *Ann. Probab.* **16** 431–478.

[26] LORETAN, M. and PHILLIPS, C. B. (1994). Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial instruments. *J. Empirical Finance* **1** 211–248.

[27] MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, Berlin.

[28] MÜLLER, U. A., DACOROGNA, M. M. and PICTET, O. V. (1996). Heavy tails in high-frequency financial data. Preprint. Olsen and Associates, Zürich.

[29] RESNICK, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.* **18** 66–138.

[30] RESNICK, S. I. (1998). Why non-linearities can ruin the heavy tailed modeller's day. In *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy-Tailed Distributions* (R. Adler, R. Feldman and M. S. Taqqu, eds.) 219–240. Birkhäuser, Boston.

[31] SAMORODNITSKY, G. and TAQQU, M. S. (1994). *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance*. Chapman and Hall, London.

[32] STAM, A. J. (1977). Regular variation in $\mathbb{R}^d_+$ and the Abel–Tauber theorem. Preprint.

[33] TAYLOR, S. J. (1986). *Modelling Financial Time Series*. Wiley, Chichester.

DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80523-1877
E-MAIL: rdavis@stat.colostate.edu

DEPARTMENT OF MATHEMATICS
P.O. BOX 800
UNIVERSITY OF GRONINGEN
NL-9700 AV GRONINGEN
THE NETHERLANDS